ARMY RESEARCH LABORATORY

# Variation in Strength of an Unconventional CH···O Hydrogen Bond in an Engineered Protein Cavity

by Rabi A. Musah, Gerard M. Jensen, Robin J. Rosenfeld,
Duncan E. McRee, David B. Goodin, and Steven W. Bunte

ARL-TR-1826

October 1998

19981020 031

DTIC QUALITY INSPECTED 4

# Abstract

We have utilized the ligand binding properties of a buried cavity created in the interior of a protein to obtain direct information about the variation in the strength of CH···O interactions between the ligand and protein. Our study shows that the strength of CH···O interactions can be modulated by over 1 kcal/mol by changes in the C–H band polarity. Consequently, several such interactions may play a significant role in the stability of macromolecular structures.

# Acknowledgments

INTENTIONALLY LEFT BLANK.

# Table of Contents

INTENTIONALLY LEFT BLANK.

# List of Figures

# List of Tables

INTENTIONALLY LEFT BLANK.

# 1. Introduction

While hydrogen bonds of the type XH···Y (X, Y = N and/or O) are essentially found in all macromolecular structures, the role of CH···O hydrogen bonds in such systems is uncertain. Recent reports [1–5] have suggested that a significant number of the CH···O contacts observed in proteins, RNA, and carbohydrates represent cohesive interactions, which may, in sufficient numbers, be significant to structure, stability, and function. The existence of short intermolecular CH···O interactions is well-established in many small-molecule crystals [6–9]. In cases where the C–H–O angle is approximately linear and the C···O bond distance is less than the combined van der Waals radii (3.3 Å), these interactions have been labeled as hydrogen bonding. Theoretical calculations have estimated the CH···O bonding interactions to be worth 1–2 kcal/mol [2, 10, 11]. However, many computational approaches utilizing force-field parameterizations do not explicitly consider such interactions as attractive, and CH···O hydrogen bonds are not generally considered in the analysis of structures. The strength of such interactions should vary considerably depending on the C–H bond polarity and may thus be different for the $C_{\alpha}H$···O interactions observed in some protein β-sheets than those for others involving the more acidic $C_{\epsilon}H$ protons of histidine [1]. While there is a correlation of the C···O distance with the C–H bond acidity for a number of organic compounds [6, 12–14], nothing is known experimentally about the strength of these interactions and how this strength varies with the polarity of the proton donor.

# 2. Background

We have utilized the ligand binding properties of a buried cavity created in the interior of a protein to obtain direct information about the variation in the strength of CH···O interactions between ligand and protein. The cavity created by the W191G mutation of cytochrome *c* peroxidase (CCP) has been shown to bind a number of cationic heterocyclic compounds [15–17]. Two such compounds, 2,3,4-trimethylthiazole (234TMT) and 3,4,5-trimethylthiazole (345TMT) bind to the

W191G cavity in essentially isosteric conformations (Figure 1).[*] Crystal structures of the protein soaked in solutions containing these compounds show clear evidence for binding in the omit electron density maps. No significant differences were observed in the structure of the protein between the ligand-free and either of the ligand-bound states. Placement of a ligand model within the omit density implies a close contact between the C5 ring proton of 234TMT and one carboxylate oxygen of Asp-235. An analogous interaction involving the C2 ring proton is observed for 345TMT. Asp-235 is observed to hydrogen-bond to other cavity-binding ligands and helps determine the cation bonding specificity of this cavity [15, 17, 18]. Thus, the absence of standard hydrogen-bond donors in 234TMT and 345TMT evidently results in the substitution of weaker alternative interactions that fulfill a similar role.

The geometries of the interactions implied by the crystal structures are consistent with a CH···O hydrogen bond [8]. The C···O distances between the ligands and the Asp-235 carboxylate oxygen were 2.95 and 2.85 Å for 234TMT and 345TMT, respectively, shorter than the combined van der Waals radii of C and O (3.3 Å). The C–H–O angles estimated from placing the geometrically optimized ligands[†] into the electron density were approximately 150° and 140° for 234TMT and 345TMT, respectively (Figure 2). The deviations of these angles from an ideal linear CH···O hydrogen bond are not outside the range (130–170°) of those observed in organic crystals [8]. Although the two compounds differ structurally only by the interchange of the N3 and C4 atoms in the thiazole ring, this difference serves to increase the acidity of the C2 proton in 345TMT

---

[*] Crystal structures of W191G soaked in 50-mM 234TMT or 345TMT were determined as previously described [17]. For each ligand, the orientation was unambiguous due to density observed for the methyl substituents and for the sulfur atom at the higher contour level. The maps were constructed by merging $F_{obsd}$ for the soaked crystal with $F_{calcd}$ from models of the W191G empty cavity. The structures of 234TMT and 345TMT bound to W191G have been submitted to the Brookhaven database (entries 1ac4 and 1ac8, respectively). Diffraction statistics for 234TMT and 345TMT, respectively, included unit cell parameters (103.76, 73.31, 44.64 Å and 106.03, 75.89, 51.13 Å), resolution (2.0 and 2.1 Å), $I/\sigma_{I(av)}$ (18.2 and 12.4), $I/\sigma_{I(last\ shell)}$ (2.00 and 1.76), number of reflections (18,219 and 19,056), percent completeness (80 and 87%), and $R_{sym}$ (0.048 and 0.062).

[†] Ligand partial charges were calculated as Kollman electrostatic potential (ESP) charges from geometry-optimized structures using density functional theory. The calculations were performed in vacuo as previously described [19] using the program Gaussian 94, Becke3LYP functional, and 6-31G* basis set. ESP partial charges were used to generate an ESP grid with the program Delphi (MSI) using a uniform dielectric constant of 1, ionic strength of 0.0, and full Coulombic boundary conditions. This potential grid was mapped onto the solvent-accessible surface of the molecule calculated with a 1.4-Å probe radius as a color spectrum from +135 kT/e (red) to +200 kT/e (blue).
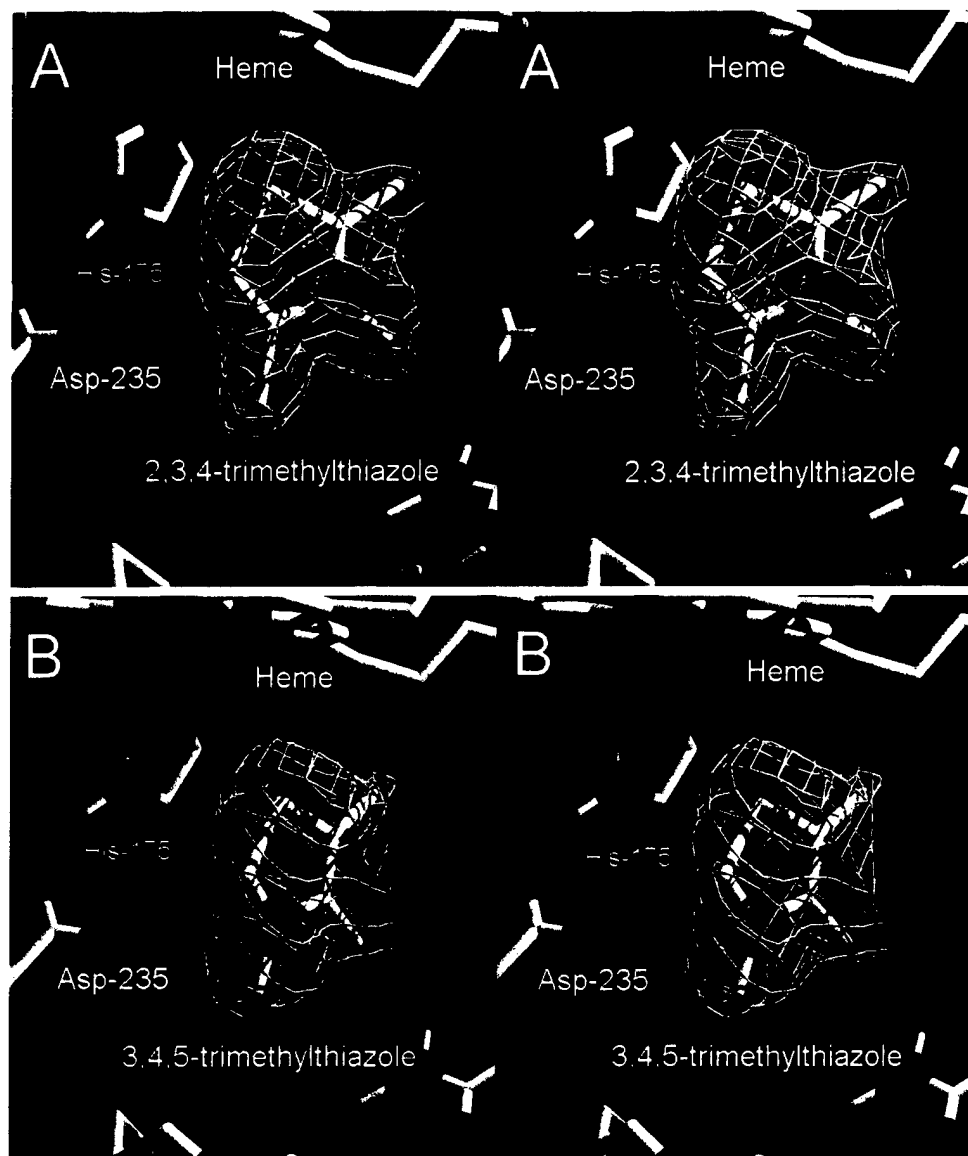
**Figure 1. Electron Density for (A) 234TMT and (B) 345TMT Binding to the W191G Cavity.**
Shown Are Stereoviews of $F_o$–$F_c$ Fourier Omit Maps Contoured at +4 σ (White) and
+9 σ (Red) Superimposed on a Model of the Ligand That Was Placed Into the Omit
Density. (See Footnote [*] on p. 2.)

($pK_a$ ≈ 17–19) [20, 21] relative to the C5 proton in 234TMT ($pK_a$ ≈ 25–30). This difference in

acidities is manifested in the electrostatic potential calculated for the two compounds and mapped

onto the solvent-accessible surface (Figure 2). Since the CH···O hydrogen bond is primarily
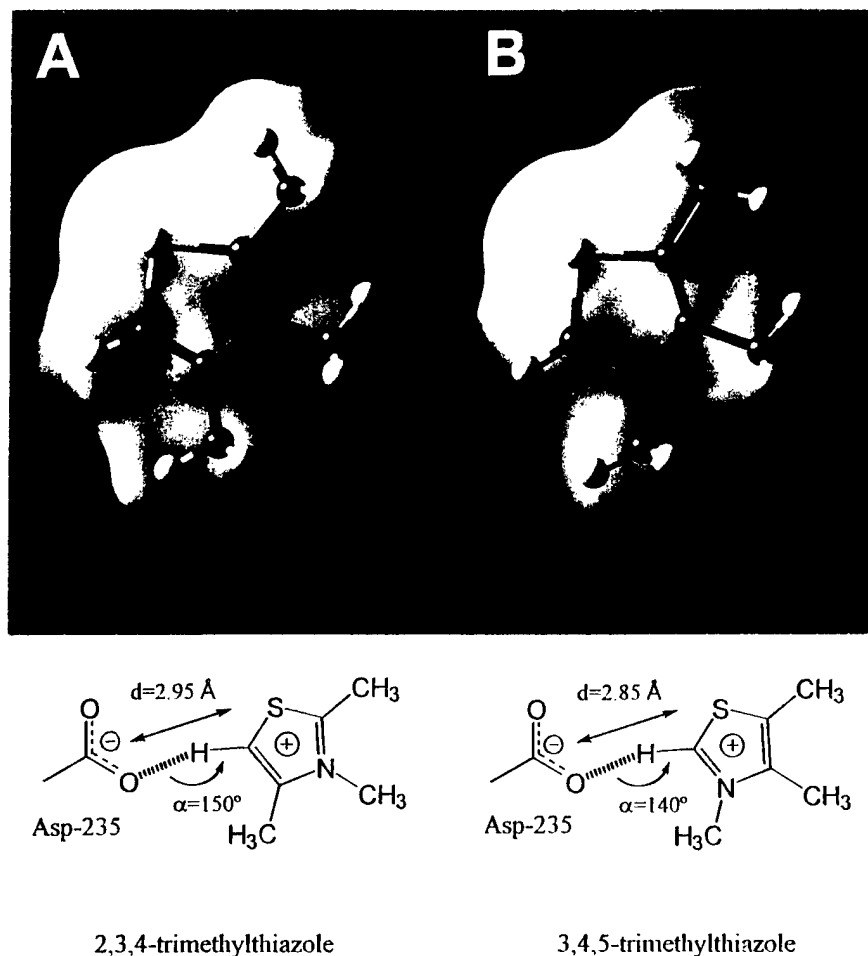
Figure 2. Electrostatic Potential Surfaces for (A) 234TMT and (B) 345TMT. The Color Map Illustrates the Localized Differences Resulting From the Increased Polarity of the C2–H Proton for 345TMT Relative to the C5–H Bond of 234TMT. (See Footnote [†] on p. 2.) At the Bottom Are the Geometries of the Interactions With Asp-235 That Are Inferred From the Placement of the Geometry-Optimized Ligand Into the Omit Electron Density.

electrostatic in nature, the calculated potentials indicate that 345TMT should form a significantly stronger interaction with Asp-235.

Thermodynamic parameters of 234TMT and 345TMT binding to W191G were determined by isothermal titration calorimetry (Table 1). The results show a five-fold decrease in $K_d$ for 345TMT relative to 234TMT, corresponding to 1.2 kcal/mol of additional binding energy for 345TMT. Due

**Table 1.** **Thermodynamic Parameters for Trimethylthiazole Binding to the Buried Cavity of W191G***

| | $K_d$ (mM) | $\Delta H$ (kcal mol$^{-1}$) | $\Delta S$ (cal mol$^{-1}$ K$^{-1}$) | $\Delta G$ (kcal mol$^{-1}$) |
|---|---|---|---|---|
| 234TMT | 1.50 (0.16) | −14 (0.6) | −33 (2) | −3.9 (0.1) |
| 345TMT | 0.20 (0.03) | −15 (0.7) | −34 (3) | −5.1 (0.2) |

to the absence of structural changes in the protein, the similar contacts made with the protein, and the fact that both compounds desolvate the cavity to the same degree, the observed difference in the free energy of binding for the two compounds should result from differences in the electrostatic interactions between ligand and protein and from differences in their desolvation energies. However, estimates of ligand desolvation energies were very similar (Table 2), indicating that the observed difference in binding free energy for the two compounds can be attributed to protein-ligand interactions.

**Table 2.** **Calculated Electrostatic Contribution to the Relative Binding Energy of 234TMT and 345TMT to the W191G Cavity***

| | $-\Delta V_{Q\mu}$ | $-\Delta V_{Q\alpha}$ | $-\Delta V_{QL}$ | $-\Delta V_B$ | $-\Delta V_{tot}$ | $-\Delta G_{solv(water)}$ |
|---|---|---|---|---|---|---|
| 234TMT | −56.88 | 3.82 | −6.44 | −7.26 | −66.76 | −26.112 |
| 345TMT | −57.83 | 2.98 | −5.26 | −7.13 | −67.24 | −26.096 |
| $\Delta\Delta G_{345TMT-234TMT}$ | −0.95 | −0.84 | 1.18 | 0.13 | −0.48 | −0.016 |

NOTE: Values are given in kcal/mol.

The electrostatic contribution to the ligand-protein interaction was estimated for the two ligands bound to the protein cavity. Partial charges were calculated for each ligand (Figure 2) to account for

---

the different C–H bond polarities, and the protein dipoles/Langevin dipoles (PDLD) method [19, 22]*

was used to account for effects of microscopic atomic polarizabilities of atoms surrounding the

ligands. The more favorable electrostatic interaction of 345TMT with the protein relative to that of

234TMT (Table 2) arises from both dipole ($\Delta V_{Q\mu}$) and induced dipole ($\Delta V_{Q\alpha}$) terms. These

interactions are partially canceled by the solvent screening effects represented in the Langevin grid

($\Delta V_{QL}$) and Born ($\Delta V_B$) terms. The net electrostatic contribution to the binding free energy ($\Delta V_{tot}$)

is approximately 0.5 kcal/mol of additional stabilization of 345TMT relative to 234TMT. This value

is in good agreement with recent *ab initio* quantum-mechanical analysis of CH···O bond energies [2].

Thus, electrostatics alone form a significant contribution to the observed difference in the binding

free energy of 1.2 kcal/mol.


# 3. Conclusion


In conclusion, the observation of untethered ligand binding to this engineered cavity shows that

the CH···O interaction is made by choice and is thus a stabilizing interaction. While it is weak,

summation of several such interactions may play a significant role in the stability of macromolecular

structures. The strength of this interaction can be modulated by over 1 kcal/mol by changes in the

C–H bond polarity. Thus, selective inclusion of polar CH···O interactions in refinement algorithms

for nuclear magnetic resonance (NMR) and crystal structures may be justified.

---

* Electrostatic calculations were performed using the PDLD method with the program POLARIS [22]. $\Delta V_{tot}$ is the sum of four terms, $\Delta V_{Q\mu}$, $\Delta V_{Q\alpha}$, $\Delta V_{QL}$, and $\Delta V_B$. $\Delta V_{Q\mu}$ is the classical electrostatic interaction of the ligand charges with all of the charges on all protein atoms. $\Delta V_{Q\alpha}$ is the interaction energy from the dipole moments induced by the electric field as a result of atomic polarizabilities. $\Delta V_{QL}$ is the energy of interaction of the ligand with the field defined by the solvent dipoles. $\Delta V_B$ is the Born energy of interaction with bulk solvent beyond the radius used to define the Langevin dipole grid. Values of $\Delta V$ have the sign of electrostatic potential, and, thus, a more positive value for $\Delta V$ corresponds to a more negative $\Delta G$. POLARIS calculations were performed with Asp-235 and Arg-48 charged, and all other amino acid residues netural. An estimate of the electrostatic interaction of the ligand charges with the solvent reaction field ($\Delta G_{solv(water)}$) for the two compounds was obtained with a continuum dielectric model (Delphi, MSI) using $\epsilon_{water} = 80$ and $\epsilon_{internal} = 2$, with a grid spacing of 0.25 Å.

# 4. References

1. Wahl, M. C., and M. Sundaralingam. *Trends Biochemical Science.* Vol. 22, pp. 97–102, 1997.

2. Ornstein, R. L., and Y.-J. Zheng. *Journal of Biomolecular Structure Dynamics.* Vol. 14, 1997.

3. Bella, J., and H. M. Berman. *Journal of Molecular Biology.* Vol. 264, pp. 734–742, 1996.

4. Derewenda, Z. S., U. Derewenda, and P. M. Kobos. *Journal of Molecular Biology.* Vol. 241, pp. 83–93, 1994.

5. Derewenda, Z. S., L. Lee, and U. Derewenda. *Journal of Molecular Biology.* Vol. 252, pp. 248–262, 1995.

6. Sutor, D. J. *Nature.* Vol. 193, pp. 68–69, 1962.

7. Hamilton, W. C., and J. A. Ibers. *Hydrogen Bonding in Solids.* New York: Benjamin, 1968.

8. Taylor, R., and O. Kennard. *Journal of American Chemical Society.* Vol. 104, pp. 5063–5070, 1982.

9. Steiner, T., and W. Saenger. *Journal of American Chemical Society.* Vol. 114, pp. 10146–10154, 1992.

10. Kollman, P., J. McKelvey, A. Johansson, and S. Rothenberg. *Journal of American Chemical Society.* Vol. 97, pp. 955–965, 1975.

11. Seiler, P., G. R. Weisman, E. D. Glendening, T. Weinhold, V. B. Johnson, and J. D. Dunitz. *Angewandte Chemie, International Edition, English.* Vol. 26, pp. 1175–1177, 1987.

12. Desiraju, G. R. *Journal of Chemical Society, Chemical Communications.* Vol. 3, pp. 179–180, 1989.

13. Desiraju, G. R. *Accounts of Chemical Research.* Vol. 29, pp. 441–449, 1996.

14. Desiraju, G. R. *Journal of Chemical Society, Chemical Communications.* Vol. 6, p. 454, 1990.

15. Fitzgerald, M. M., M. L. Trester, G. M. Jensen, D. E. McRee, and D. B. Goodin. *Protein Science.* Vol. 4, pp. 1844–1850, 1995.

16. Fitzgerald, M. M., R. A. Musah, D. E. McRee, and D. B. Goodin. *Nature and Structural Biology.* Vol. 3, pp. 626–631, 1996.

17. Fitzgerald, M. M., M. J. Churchill, D. E. McRee, and D. B. Goodin. *Biochemistry.* Vol. 33, pp. 3807–3818, 1994.

18. Miller, M. A., G. W. Han, and J. Kraut. *Proceedings of the National Academy of Science, U.S.A.* Vol. 91, pp. 11118–111122, 1994.

19. Jensen, G. M, D. B. Goodin, and S. W. Bunte. *Journal of Physical Chemistry.* Vol. 100, pp. 954–959, 1996.

20. Washabaugh, M. W., and W. P. Jencks. *Biochemistry.* Vol. 27, pp. 5044–5053, 1988.

21. Washabaugh, M. W., and W. P. Jencks. *Journal of American Chemical Society.* Vol. 111, pp. 674–683, 1989.

22. Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions.* New York: Wiley-Interscience, 1991.

| NO. OF COPIES | ORGANIZATION | NO. OF COPIES | ORGANIZATION |
|---|---|---|---|
| 2 | DEFENSE TECHNICAL INFORMATION CENTER DTIC DDA 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218 | 1 | DIRECTOR US ARMY RESEARCH LAB AMSRL D J W LYONS 2800 POWDER MILL RD ADELPHI MD 20783-1145 |
| 1 | HQDA DAMO FDQ DENNIS SCHMIDT 400 ARMY PENTAGON WASHINGTON DC 20310-0460 | 1 | DIRECTOR US ARMY RESEARCH LAB AMSRL DD J J ROCCHIO 2800 POWDER MILL RD ADELPHI MD 20783-1145 |
| 1 | OSD OUSD(A&T)/ODDDR&E(R) R J TREW THE PENTAGON WASHINGTON DC 20301-7100 | 1 | DIRECTOR US ARMY RESEARCH LAB AMSRL CS AL TA 2800 POWDER MILL RD ADELPHI MD 20783-1145 |
| 1 | DPTY CG FOR RDE HQ US ARMY MATCOM AMCRD MG BEAUCHAMP 5001 EISENHOWER AVE ALEXANDRIA VA 22333-0001 | 3 | DIRECTOR US ARMY RESEARCH LAB AMSRL CI LL 2800 POWDER MILL RD ADELPHI MD 20783-1145 |
| 1 | INST FOR ADVNCD TCHNLGY THE UNIV OF TEXAS AT AUSTIN PO BOX 202797 AUSTIN TX 78720-2797 | | ABERDEEN PROVING GROUND |
| 1 | DARPA B KASPAR 3701 N FAIRFAX DR ARLINGTON VA 22203-1714 | 4 | DIR USARL AMSRL CI LP (305) |
| 1 | NAVAL SURFACE WARFARE CTR CODE B07 J PENNELLA 17320 DAHLGREN RD BLDG 1470 RM 1101 DAHLGREN VA 22448-5100 | | |
| 1 | US MILITARY ACADEMY MATH SCI CTR OF EXCELLENCE DEPT OF MATHEMATICAL SCI MAJ M D PHILLIPS THAYER HALL WEST POINT NY 10996-1786 | | |

INTENTIONALLY LEFT BLANK.

# REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project(0704-0188), Washington, DC 20503.

| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | October 1998 | Final, Jan 96 - Jan 97 |

| 4. TITLE AND SUBTITLE | 5. FUNDING NUMBERS |
|---|---|
| Variation in Strength of an Unconventional CH···O Hydrogen Bond in an Engineered Protein Cavity | 1L161102AH43 |

**6. AUTHOR(S)**

Rabi A. Musah,* Gerard M. Jensen,* Robin J. Rosenfeld,* Duncan E. McRee,* David B. Goodin,* and Steven W. Bunte

| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|
| U.S. Army Research Laboratory<br>ATTN: AMSRL-WM-BD<br>Aberdeen Proving Ground, MD 21005-5066 | ARL-TR-1826 |

| 9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES) | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER |
|---|---|
| | |

**11. SUPPLEMENTARY NOTES**

*Department of Molecular Biology, MB8, The Scripps Reseach Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037

| 12a. DISTRIBUTION/AVAILABILITY STATEMENT | 12b. DISTRIBUTION CODE |
|---|---|
| Approved for public release; distribution is unlimited. | |

**13. ABSTRACT (Maximum 200 words)**

We have utilized the ligand binding properties of a buried cavity created in the interior of a protein to obtain direct information about the variation in the strength of CH···O interactions between the ligand and protein. Our study shows that the strength of CH···O interactions can be modulated by over 1 kcal/mol by changes in the C–H band polarity. Consequently, several such interactions may play a significant role in the stability of macromolecular structures.

| 14. SUBJECT TERMS | | | 15. NUMBER OF PAGES |
|---|---|---|---|
| hydrogen bond, ab initio, calculations, electrostatic modeling | | | 15 |
| | | | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UL |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18     298-102

INTENTIONALLY LEFT BLANK.

# USER EVALUATION SHEET/CHANGE OF ADDRESS

This Laboratory undertakes a continuing effort to improve the quality of the reports it publishes. Your comments/answers to the items/questions below will aid us in our efforts.

1. ARL Report Number/Author _____ARL-TR-1826 (Musah)_____ Date of Report __October 1998__

2. Date Report Received _____

3. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which the report will be used.) _____

_____

_____

4. Specifically, how is the report being used? (Information source, design data, procedure, source of ideas, etc.) _____

_____

_____

5. Has the information in this report led to any quantitative savings as far as man-hours or dollars saved, operating costs avoided, or efficiencies achieved, etc? If so, please elaborate. _____

_____

_____

6. General Comments. What do you think should be changed to improve future reports? (Indicate changes to organization, technical content, format, etc.) _____

_____

_____

_____

| | |
|---|---|
| | _____ |
| | Organization |
| CURRENT ADDRESS | Name                        E-mail Name |
| | Street or P.O. Box No. |
| | City, State, Zip Code |

7. If indicating a Change of Address or Address Correction, please provide the Current or Correct address above and the Old or Incorrect address below.

| | |
|---|---|
| | _____ |
| | Organization |
| OLD ADDRESS | Name |
| | Street or P.O. Box No. |
| | City, State, Zip Code |

(Remove this sheet, fold as indicated, tape closed, and mail.)
**(DO NOT STAPLE)**
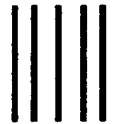
**DEPARTMENT OF THE ARMY**

OFFICIAL BUSINESS

## BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO 0001,APG,MD

POSTAGE WILL BE PAID BY ADDRESSEE

DIRECTOR
US ARMY RESEARCH LABORATORY
ATTN AMSRL WM BD
ABERDEEN PROVING GROUND MD 21005-5066